

The listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1.-41. (Canceled)

42. (New) A speech synthesis device, comprising:

voice unit storage means for storing a plurality of pieces of voice unit data representing voice units;

phoneme storage means for storing a plurality of pieces of phoneme data each of which is a phoneme or comprises phoneme fragments composing a phoneme;

cadence prediction means for inputting sentence information representing a sentence to predict the cadence of voice units composing the sentence;

selecting means for selecting voice unit data satisfying predetermined conditions out of the plurality of pieces of voice unit data stored in the voice unit storage means, wherein the predetermined conditions are that the voice unit data to be selected matches in its reading with the voice unit composing the sentence and has a correlation greater than a predetermined amount with a cadence prediction result by the cadence prediction means;

missing part cadence prediction means for predicting the cadence of voice units which have been decided not to satisfy the predetermined conditions by the selection means;

missing part synthesis means for specifying phonemes contained in the voice unit decided not to satisfy the predetermined condition by the selection means out of the voice units composing the sentence, for acquiring phoneme data representing the specified phoneme or phoneme fragments composing the specified phoneme from the phoneme storage means, for converting the acquired phoneme data so that the

phoneme or phoneme fragments represented by the acquired phoneme data matches with a cadence prediction result by the missing part cadence prediction means, and for interconnecting the converted data, thereby synthesizing speech data representing a waveform of the voice unit; and

creation means for interconnecting the voice unit data selected by the selection means and the speech data synthesized by the missing part synthesis means, thereby creating data representing synthesis speech.

43. (New) The speech synthesis device according to claim 42, wherein the selection means selects the voice unit data out of the plurality of pieces of voice unit data stored in the voice unit storage means under the predetermined conditions further including that the presence or absence of nasalization or devocalization of the voice unit data matches with the cadence prediction result by the cadence prediction means.

44. (New) The speech synthesis device according to claim 43, wherein the device further comprises utterance speed conversion means for acquiring utterance speed data specifying conditions of a speed for producing the synthesis speech created by the reaction means and for converting the voice unit data and/or speech data so as to represent a speech to be produced at a speed satisfying the conditions specified by the utterance speed data.

45. (New) The speech synthesis device according to claim 44, wherein the utterance speed conversion means operates to convert the voice unit data and/or the speech data so as to represent a speech to be uttered at a speed to be produced at a speed satisfying the conditions specified by the utterance speed data, by eliminating a segment representing a phoneme fragment from voice unit data and/or speech data composing data representing the synthesis speech or by adding a segment representing a phoneme fragment to the voice unit data and/or speech data.

46. (New) The speech synthesis device according to any of claims 42, wherein the voice unit storage means operates to associate phonetic data representing a reading of voice unit with the voice unit data, and the selection means operates to handle voice unit data which is associated with phonetic data representing a reading matching with the reading of the voice unit composing the sentence, as voice unit whose reading is common with the voice unit.

47. (New) The speech synthesis device according to any of claims 43, wherein the voice unit storage means operates to associate phonetic data representing a reading of voice unit with the voice unit data, and the selection means operates to handle voice unit data which is associated with phonetic data representing a reading matching with the reading of the voice unit composing the sentence, as voice unit whose reading is common with the voice unit.

48. (New) The speech synthesis device according to any of claims 44, wherein the voice unit storage means operates to associate phonetic data representing a reading of voice unit with the voice unit data, and the selection means operates to handle voice unit data which is associated with phonetic data representing a reading matching with the reading of the voice unit composing the sentence, as voice unit whose reading is common with the voice unit.

49. (New) The speech synthesis device according to any of claims 45, wherein the voice unit storage means operates to associate phonetic data representing a reading of voice unit with the voice unit data, and the selection means operates to handle voice unit data which is associated with phonetic data representing a reading matching with the reading of the voice unit composing the sentence, as voice unit whose reading is common with the voice unit.

50. (New) A speech synthesis method performed by a speech synthesis device having storage means and processing means, the method comprising the steps of:

storing in the storage means a plurality of pieces of voice unit data representing voice units;

storing in the storage means a plurality of pieces of phoneme data each of which is a phoneme or comprises phoneme fragments composing a phoneme;

inputting in the processing means sentence information representing a sentence to predict the cadence of voice units composing the sentence;

selecting, in the processing means, voice units satisfying predetermined conditions out of the plurality of pieces of voice unit data stored in the storage means, wherein the predetermined conditions are that the voice unit data to be selected matches in its reading with the voice unit composing the sentence and has a correlation greater than a predetermined amount with a cadence prediction result;

predicting in the processing means the cadence of voice units which have been decided not to satisfy the predetermined conditions;

in the processing means, specifying phonemes contained in the voice unit decided not to satisfy the predetermined conditions out of the voice units composing the sentence, acquiring phoneme data representing the specified phoneme or phoneme fragments composing the specified phoneme from the storage means, converting the acquired phoneme data so that the phoneme or phoneme fragments represented by the acquired phoneme data matches with a cadence prediction result, and interconnecting the converted data, thereby synthesizing speech data representing a waveform of the voice unit; and

in the processing means, interconnecting the selected voice unit data and the synthesis speed data, thereby creating data representing synthesis speech.

51. (New) The speech synthesis method according to claim 50, wherein the processing means operates to select the voice unit data out of the plurality of pieces of voice unit data stored in the storage means under the predetermined conditions further including that the presence or absence of nasalization or devocalization of the voice unit data matches with the cadence prediction result.

52. (New) A computer readable medium which records a computer program causing a computer to operate as:

voice unit storage means for storing a plurality of pieces of voice unit data representing voice units;

phoneme storage means for storing a plurality of pieces of phoneme data each of which is a phoneme or comprises phoneme fragments composing a phoneme;

cadence prediction means for inputting sentence information representing a sentence to predict the cadence of voice units comprising the sentence;

selecting means for selecting voice unit data satisfying predetermined conditions out of the plurality of pieces of voice unit data stored in the voice unit storage means, wherein the predetermined conditions are that the voice unit data to be selected matches in its reading with the voice unit composing the sentence and has a correlation greater than a predetermined amount with a cadence prediction result by the cadence prediction means;

missing part cadence prediction means for predicting the cadence of voice units which have been decided not to satisfy the predetermined conditions by the selection means;

missing part synthesis means for specifying phonemes contained in the voice unit decided not to satisfy the predetermined condition by the selection means out of the voice units composing the sentence, for acquiring phoneme data representing the specified phoneme or phoneme fragments composing the specified phoneme from the phoneme storage means, for converting the acquired phoneme data so that the

phoneme or phoneme fragments represented by the acquired phoneme data matches with a cadence prediction result by the missing part cadence prediction means, and for interconnecting the converted data, thereby synthesizing speech data representing a waveform of the voice unit; and

creation means for interconnecting the voice unit data selected by the selection means and the speech data synthesized by the missing part synthesis means, thereby creating data representing synthesis speech.

53. (New) The computer readable medium according to claim 52, wherein the selection means selects the voice unit data out of the plurality of pieces of voice unit data stored in the voice unit storage means under the predetermined conditions further including that the presence or absence of nasalization or devocalization of the voice unit data matches with the cadence prediction result by the cadence prediction means.